DOCUMENT RESUME

ED 157 445 HE 010 275

AUTHOR	Centra, John A.
TITLE	How Universities Evaluate Faculty Performance: A Survey of Department Heads.
INSTITUTION	Graduate Record Examinations Board, Princeton, N.J.
REPORT NO	GREB-75-5bR
PUB DATE	Jul 77
NOTE	26p.; Not available in hard copy due to marginal legibility of original
AVAILABLE FROM	Graduate Record Examination Program, Educational Testing Service, Princeton, New Jersey 08540 ($1.50)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*College Faculty; *Department Directors (School); Departments; Effective Teaching; *Evaluation Criteria; *Faculty Evaluation; Higher Education; Institutional Characteristics; *National Surveys; Questionnaires; Research; Student Evaluation of Teacher Performance; *Teacher Evaluation
IDENTIFIERS	*Department Chairpersons

ABSTRACT

Department heads from 134 institutions (mainly universities) indicated the weight they generally give to various criteria for evaluating individual faculty members. The questionnaire they responded to included: criteria used for evaluating overall faculty performance; sources of information for evaluating teaching; and kinds of information used for judging scholarship or research performance. Although teaching, research, and service are generally acknowledged as the major functions of most universities, responses by the 453 department heads indicated that public or university service is usually given little importance in evaluating faculty members for decisions regarding tenure, salary, and promotion. The research universities emphasize research and scholarship. The so-called doctoral-granting universities and the comprehensive universities and colleges said that teaching was ranked first in importance, followed closely by research. The results of the study also suggest an increase in the use of some of the more systematic and tangible kinds of evidence for evaluating teaching performance (e.g., student ratings), though the use of such evidence falls short of what department heads think is needed. In general, the evaluation of research and scholarship depends very much on the level of the institution and the type of department. For example, peer judgments of research and the number of articles in quality journals are much more important in social science departments in the research universities than in the same departments in the comprehensive universities and colleges. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED157445

THE GRADUATE RECORD EXAMINATIONS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ETS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

GRE

How Universities Evaluate Faculty Performance: A Survey of Department Heads

John A. Centra

GRE Board Research Report GREB No. 75-5bR

July 1977

This report presents the findings of a research project funded by and carried out under the auspices of the Graduate Record Examinations Board.

EDUCATIONAL TESTING SERVICE, PRINCETON, NEW JERSEY □ BERKELEY, CALIFORNIA □ EVANSTON, ILLINOIS

HE 010 275

Abstract

Department heads from 134 institutions (mainly universities) indicated the weight they generally give to various criteria for evaluating individual faculty members. The questionnaire they responded to included: (a) the criteria used for evaluating overall faculty performance; (b) the sources of information for evaluating teaching; and (c) the kinds of information used for judging scholarship or research performance. Although teaching, research, and service are generally acknowledged as the major functions of most universities, responses by the 453 department heads indicated that public or university service is usually given little importance in evaluating faculty members for decisions regarding tenure, salary, and promotion. The Research Universities emphasize research and scholarship, as might be expected. The so-called Doctoral-Granting Universities and the Comprehensive Universities and Colleges said that teaching was ranked first in importance, followed closely by research. The results of the study also suggest an increase in the use of some of the more systematic and tangible kinds of evidence for evaluating teaching performance (e.g., student ratings), though the use of such evidence falls far short of what department heads think is needed. In general, the evaluation of research and scholarship depends very much on the level of the institution and the type of department. For example, peer judgments of research and the number of articles in quality journals are much more important in social science departments in the Research Universities than in the same departments in the Comprehensive Universities and Colleges.

Additional copies of this report, at $1.50 per copy, may be ordered from:

Graduate Record Examinations Program
Educational Testing Service
Princeton, NJ 08540

Copyright © 1977 by Educational Testing Service. All rights reserved.

How Universities Evaluate Faculty Performance: A Survey of Department Heads

GREB No. 75-5bR

Decisions concerning tenure and promotion are becoming increasingly more difficult to make in colleges and universities. During the expansion years of the sixties, most departments were concerned largely with recruiting and keeping competent faculty members; but enrollments have tapered off and departments are no longer adding staff members. This steady-state condition has put many institutions at or near prescribed limits in the percentage of faculty on tenure; in fact, the American Association of University Professors' *Annual Report on the Economic Status of the Profession* found that 60 percent of all faculty members had tenure in 1976.

In contrast to the last decade, institutions are now often forced to make fine distinctions between generally competent staff members. What information do departments use in making these decisions? How do the critical decision makers, such as department heads, think an individual faculty member's performance ought to be evaluated? These questions were investigated in this study by surveying practices at a large number of universities.

Previous Research

Most studies of how faculty performance is evaluated have been based on reports from academic deans or vice presidents. One of the earliest studies was by Gustad (1961) and included a national sample of colleges and universities. Astin and Lee (1967) repeated the survey about five years later with a similarly extensive sample of all types of postsecondary institutions. Both studies asked administrators to indicate the importance of various sources of information in evaluating, first, faculty performance and, second, teaching performance. The results were analyzed by type of institution. In both surveys the university deans of arts and sciences ranked department heads as the most important source of information. Although the deans ranked informal student opinion and systematic student ratings high in 1961, five years later they indicated that students were less a factor in promotions than they had been. Over 90 percent of the arts and science deans in the 110 universities in the Astin and Lee survey indicated that classroom teaching and research were given equally high weight in assessing faculty performance.

Seldin (1975) repeated the Gustad and the Astin and Lee surveys in 1974 with academic deans in liberal arts colleges in order to examine changes that might have taken place in the eight-year period. The deans reported more emphasis on systematic student ratings in evaluating teaching (an increase from 11 to 29 percent of the colleges) and slightly less emphasis on research in evaluating overall faculty performance. The increased use of student ratings in making decisions on faculty advancement was also reported in studies that included doctoral-level universities (Bejar, 1975; Boyd & Schietinger, 1976). A Southern Regional Education Board study found, in addition, that the major purpose of evaluation in doctoral-level universities was to make decisions on faculty advancement, rather than to improve instruction, and that department heads had greater responsibility than academic deans or vice presidents in making these decisions (Boyd & Schietinger, 1976).

One of the few studies that sampled faculty perceptions of evaluation procedures found that the most influential factors in promotions and tenure decisions were publications, the department head's evaluation, and student ratings (Thorne, Scott, & Beaird, 1976). This study, which was conducted at institutions in the Oregon State system, also reported that grant support was considered to be influential in the physical science departments.

Although the past studies point to the importance of department heads in making decisions on advancement, none surveyed these individuals' perceptions of current practices. With the exception of the Oregon State study, the earlier studies had also investigated administrators' perceptions of institutionwide procedures rather than the criteria used at the department level. Faculty performance might be viewed quite differently in different fields of study. In particular, the evaluation of research and scholarship, which the prior studies had not looked at in detail, might be expected to vary considerably by department.

Purpose of the Present Study

The present study reports the importance that department heads or chairmen give to various criteria in evaluating individual faculty members.[1] In addition to overall faculty performance, the study discusses the chairmen's views of the use and importance of specific criteria for assessing both teaching and research performance. To determine the ex-

[1]Both titles—department head and chairman—apply to the groups surveyed and are used in this report although it is recognized that the terms are not always synonymous. Use of *chairman* rather than *chairperson* or *chair* follows current APA style guidelines, as published in *American Psychologist*, June 1977, p. 492.

tent to which current practices coincide with preferred practices, respondents also were asked to indicate what importance they thought *should* be attached to each criterion. Comparisons are discussed for four departmental subgroups and for three different types of universities.

Procedure

The questionnaire used in the study (see Appendix A) partially overlapped the forms first used in the Gustad (1961) and Astin and Lee (1967) studies and later adapted by Seldin (1975). It differed from the forms used in the previous studies in that (1) a question regarding the criteria used to evaluate scholarship or research performance was added, and (2) the criteria used to evaluate overall faculty or teaching performance were somewhat altered or extended. Four copies of the questionnaire were sent to the graduate deans of 168 universities (a 50 percent random sample of the members of the Council of Graduate Schools in the United States). The deans were asked to distribute the questionnaires to four heads of departments from a variety of disciplines, giving preference to departments that offered both graduate and undergraduate degree programs. One or more usable questionnaires were received from 134 universities—a participation rate of about 80 percent. Questionnaire returns from department heads, based on four copies per university, amounted to about 67 percent from all universities that were contacted, or 85 percent of the 134 universities known to have distributed the questionnaires. Eight graduate deans said they could not participate in the study. A total of 453 usable questionnaires were included in the data analyses.

For purposes of analysis, the universities were grouped according to a classification scheme developed by the Carnegie Commission on Higher Education (1973). Three groups were formed by combining Carnegie subcategories. The first level, Research Universities, included (a) the 100 universities that led in federal financial support in 1969–70 and that awarded at least 50 doctor's degrees in 1969–70 or (b) the 50 institutions that led in terms of total number of doctor's degrees awarded during the years 1960 to 1970. In all cases, Ph.D.'s, Ed.D.'s, M.D.'s on the same campus, and other doctor's degrees were counted. There were 158 departments from these Research Universities in the study.

The second level, Doctoral-Granting Universities, combined two categories. The first category consisted of institutions that awarded 40 or more doctoral degrees in 1969–70 or received at least $3 million in total federal financial support in either 1969–70 or 1970–71. The second category included institutions that awarded at least 10 doctoral degrees in 1969–70 and a few new doctoral-granting institutions that may be expected to increase the number of degrees awarded within a few years. There were 122 departments from these Doctoral-Granting Universities in the study, 80 percent of which were in the first category.

The third level, Comprehensive Universities and Colleges, consisted mainly of institutions that offered a liberal arts program and at least two professional or occupational programs. Many of them offered a master's degree, but lacked a doctoral program or had an extremely limited doctoral program. About 80 percent of the 173 departments in this group were at institutions that fit this description. The remaining 20 percent were at institutions that offered a liberal arts program and at least one professional/occupational program or were specialized institutions (for example, schools of engineering or technology).

Of the 453 departments in the study, about two-thirds were at public institutions. Sixty-four percent of these institutions offered a Ph.D. or other doctoral degree, and 36 percent offered a master's as their highest degree. Only 3 of the 453 departments offered nothing higher than a bachelor's degree.

The questionnaire respondents represented about 80 different academic disciplines. To evaluate possible differences in the responses of department heads from different disciplines, the departments were grouped into four categories: humanities, social sciences, natural sciences, and professional fields. Humanities included all languages and literature, philosophy, religion, speech, fine arts, music, and dramatic arts. Social sciences included geography, anthropology, economics, history, government, psychology, and sociology. Natural sciences included all of the biological sciences, chemistry, physics, geology, and mathematics. Professional fields included mainly departments of education, engineering, and business. In addition, a total of about a dozen responses came from department heads in fields such as architecture, agriculture, forestry, home economics, veterinary medicine, and law.

Table 1 includes the number of departments in each group and at each university level. There were 102 departments in the social sciences, 104 in the humanities, 131 in the natural sciences, and 116 in the professional fields. For each of the three levels of universities determined by the Carnegie classification scheme, there were between 26 and 47 departments in each of the four discipline groups, thus allowing department group by university level interactions to be investigated.

Table 1. Number of Departments in the Study within Each Department Group and at Each University Level

University Levels	Department Groups				Total
	Professional Fields	Social Sciences	Humanities	Natural Sciences	
Research Universities	43	36	36	43	158
Doctoral-Granting Universities	26	27	27	42	122
Comprehensive Universities and Colleges	47	39	41	46	173
Total	116	102	104	131	453

Each criterion in the questionnaire was responded to on a five-point scale as follows: not available, not a factor, minor factor, major factor, and extremely critical factor. The average response for each criterion was computed for each de-

partment group and university level. Analysis of variance and canonical discriminant function analysis were used in the study to investigate response differences.

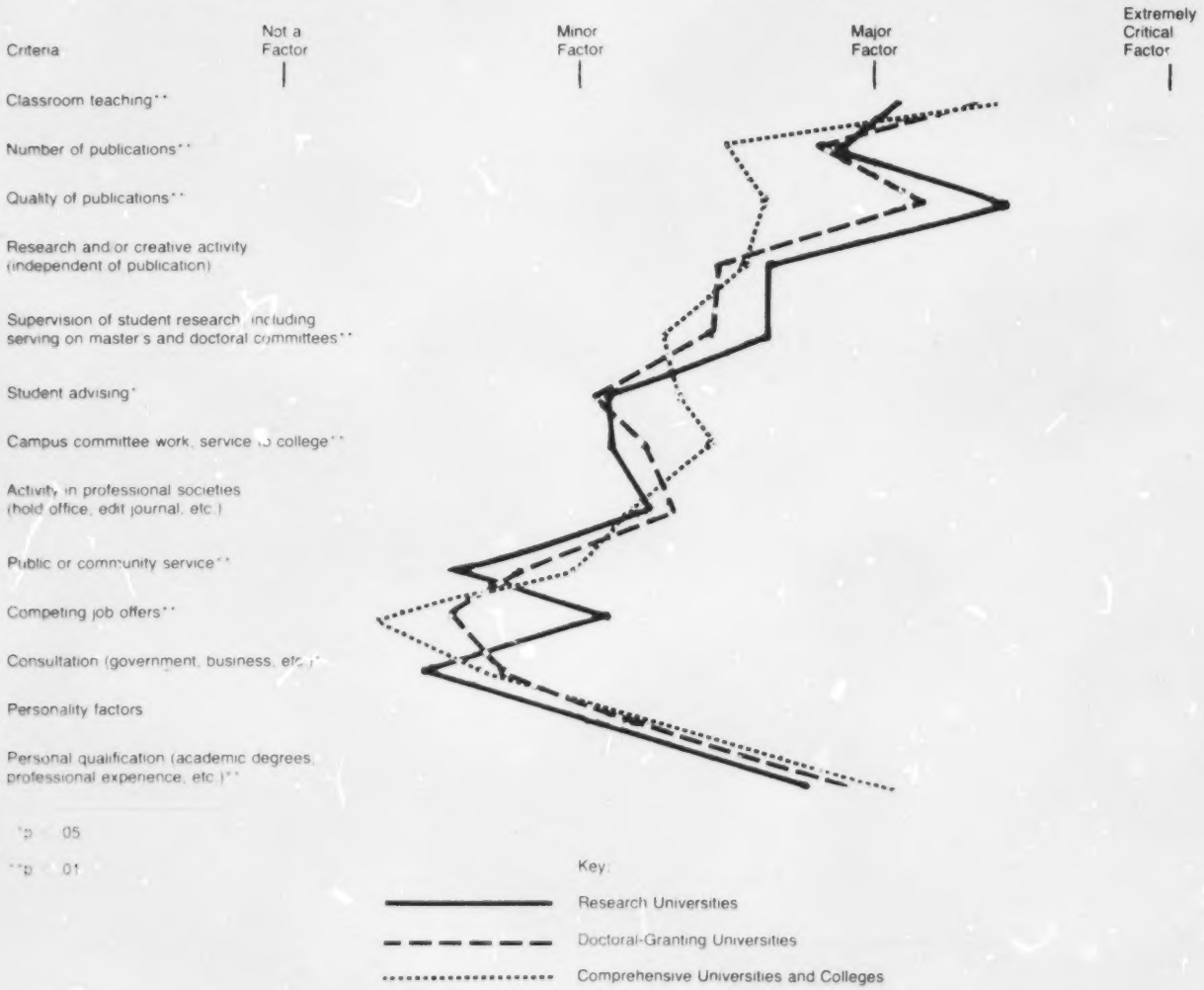
Results

The results are presented in three sections: the criteria used for evaluating overall faculty performance, the sources of information for evaluating teaching, and the kinds of information used for judging scholarship or research performance. In each section the data were analyzed by level of university and by department subgroup. Although differences for department groups within each of the university levels (department by level interaction) were also investigated, these interactions were significant only in the evaluation of scholarship or research performance; canonical discriminant function analysis was used to illustrate this interaction.

Criteria in Assessing Total Faculty Performance

The questionnaire included 13 general criteria that might be used in evaluating faculty members. These are listed in Figures 1 and 2 and in Tables 2, 3, and 4, along with the summaries of the responses of department chairmen. Classroom teaching, quality of publications, and personal qualifications (academic degrees and professional experience) were most frequently reported as major or extremely critical

Figure 1. Average weight given to various criteria for evaluating total faculty performance, by university level



factors in judging faculty members. As indicated in Figure 1, however, department chairmen at the three levels of institutions gave different emphases to each of the three criteria. Less weight was placed on classroom teaching or personal qualifications in the Research Universities than in Doctoral-Granting Universities or Comprehensive Universities and Colleges. On the other hand, the reverse was true for quality of publications, which, not surprisingly, was judged more important in the Research Universities than at either of the other two university levels. Chairmen in Comprehensive Universities and Colleges not only placed less emphasis on quality of publications, but also gave it essentially the same weight as the number of publications and as college service (for example, campus committee work). Other criteria in which universities differed according to Carnegie classification were the following: supervision of

Figure 2. Average weight given to various criteria for evaluating total faculty performance, by departmental subgroup

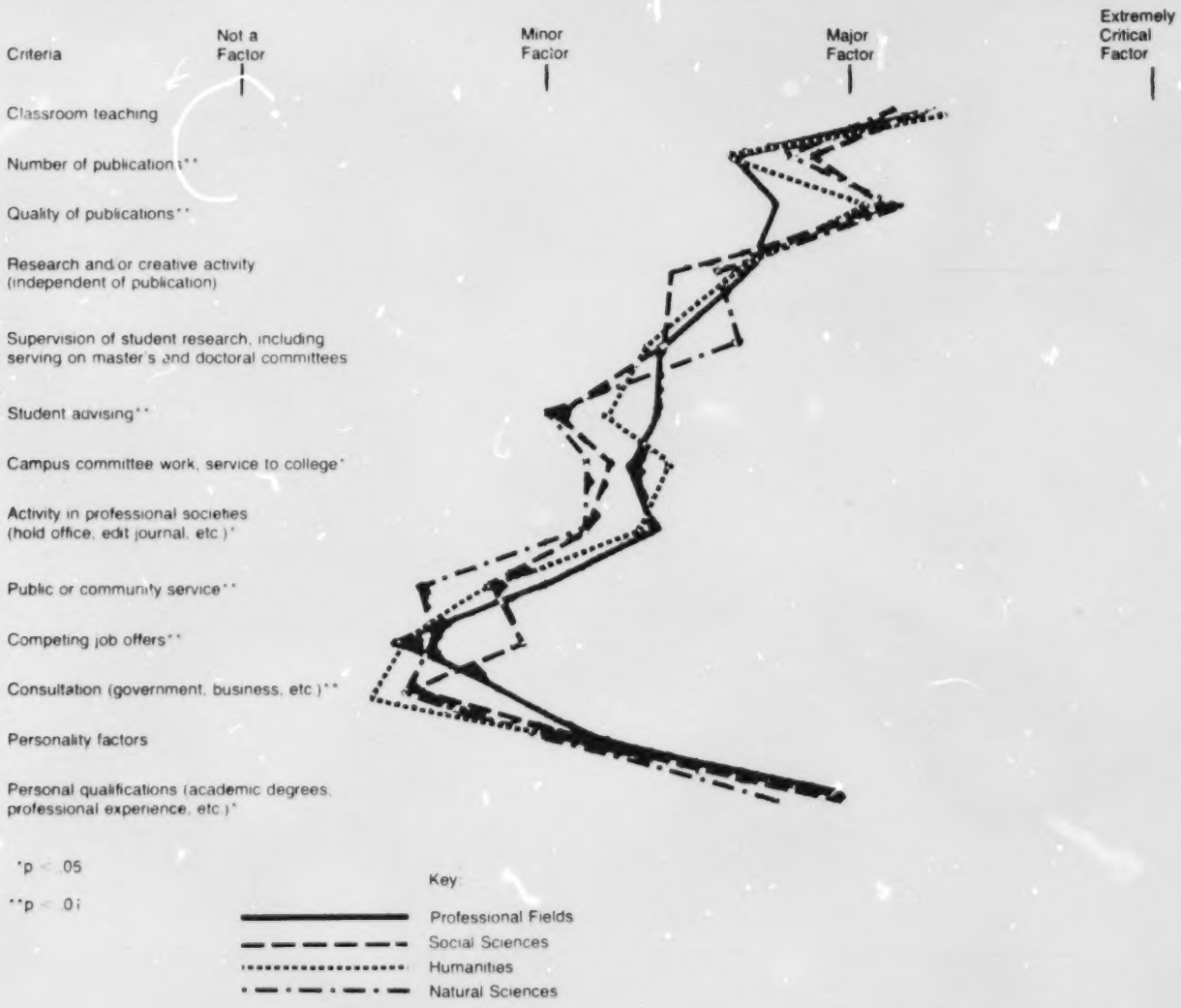


Table 2. Current and Preferred Criteria for Evaluating Total Faculty Performance

	Percentage Indicating Extremely Critical Factor or Not a Factor															
	Current Use and Importance in Personnel Recommendations								Importance Each *Should* Have in Personnel Recommendations							
	Professional Fields N=116		Social Sciences N=102		Humanities N=104		Natural Sciences N=131		Professional Fields N=116		Social Sciences N=102		Humanities N=104		Natural Sciences N=131	
	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.
Classroom teaching	38	0	35	1	44	1	31	0	59	0	46	1	69	0	44	0
Number of publications	10	2	15	2	13	6	12	2	4	1	8	1	6	7	9	1
Quality of publications	20	9	41	4	38	3	36	2	34	0	55	0	52	1	49	1
Research and/or creative activity (independent of publication)	12	3	8	7	17	6	18	2	20	1	15	3	23	0	21	1
Supervision of student research, including serving on master's and doctoral committees	6	8	5	9	8	8	6	5	16	2	8	6	15	2	15	2
Student advising	8	11	2	24	6	17	3	17	13	1	10	8	11	7	6	5
Campus committee work, service to college	2	5	3	9	5	4	2	9	5	3	1	6	4	3	2	5
Activity in professional societies (hold office, edit journal, etc.)	2	8	2	10	4	5	2	8	3	3	5	5	6	3	4	4
Public or community service	2	24	2	29	2	36	1	47	3	16	3	25	2	27	0	36
Competing job offers	1	53	7	31	4	47	4	44	1	58	3	38	2	43	2	49
Consultation (government, business, etc.)	1	23	1	45	1	47	0	50	2	21	2	35	0	43	1	40
Personality factors	5	15	2	25	10	21	4	18	9	16	1	33	7	25	2	21
Personal qualifications (academic degrees, professional experience, etc.)	27	3	32	7	28	6	23	8	23	3	30	6	32	5	21	8

student research (greatest emphasis in Research Universities), student advising (greatest emphasis in Comprehensive Universities and Colleges), and competing job offers (most important in Research Universities).

Figure 2 illustrates the differences in emphasis on each criterion according to departmental subgroup. Publications (quality and number) were more important in the natural and social sciences than in the professional schools and the humanities. Student advising, public service, and consultation were given more weight in the professional departments than in the other three department groups.

Current and preferred criteria for evaluating faculty members for decisions on promotions are given in Tables 2 and 3 (by departments) and in Table 4 (by university level). The absolute differences between current and preferred emphases were also investigated and have been taken into account in the following discussion. The percentages of department heads reporting each criterion as an "extremely critical factor" or "not a factor" are reported in Table 2. As this table indicates, there are not only variations between the department subgroups but, to some extent, within the subgroups as well. For example, 41 percent of the social science department heads reported quality of publications to be an extremely critical factor and 4 percent reported it *not* to be a factor. In general Table 2 indicates that quality of publications was a little more frequently chosen as a critical factor than teaching was in the social sciences and natural sciences, while the reverse was true in the professional fields and the humanities. Criteria that chairmen preferred to emphasize more were quality of publications, supervision of student research, advising students, and, to some extent, unpublished research or creative activity (Tables 2 and 3). The number of publications should, according to the chairmen, be a less critical factor. Several criteria were not

Table 3. Ranking of Current and Preferred Criteria for Evaluating Faculty for Promotion, Salary Increase, or Tenure, by Departmental Subgroup[1]

	Current Use and Importance				Importance Each *Should* Have			
	P	S	H	N	P	S	H	N
Classroom teaching	1	1	1	1.5	1	2	1	2
Number of publications	5	4	5	3	7	5	7	6
Quality of publications	3	2	2	1.5	2	1	2	1
Research and/or creative activity (independent of publication)	4	6	4	6	3	4	4	4
Supervision of student research, including serving on master's and doctoral committees	6	5	7	5	5	6	5	3
Student advising	7	9.5	9	10	6	7	6	7
Campus committee work, service to college	9	7	6	8	9	8	8	9
Activity in professional societies (hold office, edit journal, etc.)	8	8	8	7	8	9	9	8
Public or community service	11	12	11	12	11	10.5	11	11.5
Competing job offers	13	11	12	11	13	12	12	13
Consultation (government, business, etc.)	12	13	13	13	12	13	13	11.5
Personality factors	10	9.5	10	9	10	10.5	10	10
Personal qualifications (academic degrees, professional experience, etc.)	2	3	3	4	4	3	3	5

[1]P = Professional Fields; N = 116 (rho = .97)
S = Social Sciences; N = 102 (rho = .94)
H = Humanities; N = 104 (rho = .94)
N = Natural Sciences; N = 131 (rho = .90)

Rank difference correlations between current importance and preferred importance are shown in parentheses.

Table 4. Ranking of Current and Preferred Criteria for Evaluating Faculty for Promotion, Salary Increase, or Tenure, by University Level[1]

	Current Use and Importance				Importance Each *Should* Have			
	All	I	II	III	All	I	II	III
Classroom teaching	1	2	1	1	1	2	1	1
Number of publications	4	3	4	5	6	6	6	7
Quality of publications	2	1	2	3	2	1	2	2
Research and/or creative activity (independent of publication)	5	5	5	4	4	3	4	4
Supervision of student research, including serving on master's and doctoral committees	6	6	6	8	5	4	5	6
Student advising	9	9	10	7	7	7	7	5
Campus committee work, service to college	7	9	8	6	9	9	9	8
Activity in professional societies (hold office, edit journal, etc.)	8	7	7	9.5	8	8	8	9
Public or community service	11	12	11	11	11	12	12	11
Competing job offers	12	9	13	13	13	11	13	13
Consultation (government, business, etc.)	13	13	12	12	12	13	11	12
Personality factors	10	11	9	9.5	10	10	10	10
Personal qualifications (academic degrees, professional experience, etc.)	3	4	3	2	3	5	3	3

[1]Departments grouped by university classification as follows:
I = Research Universities; N = 158 (rho = .92)
II = Doctoral-Granting Universities; N = 122 (rho = .95)
III = Comprehensive Universities and Colleges; N = 173 (rho = .95)

Rank difference correlations between current importance and preferred importance are shown in parentheses.

factors in evaluating faculty performance in many departments—in particular, competing job offers, consultation, public or community service, and personality factors. Moreover, none of these deserved more weight, according to the respondents. The respondents also were generally satisfied with the great importance given to personal qualifications.

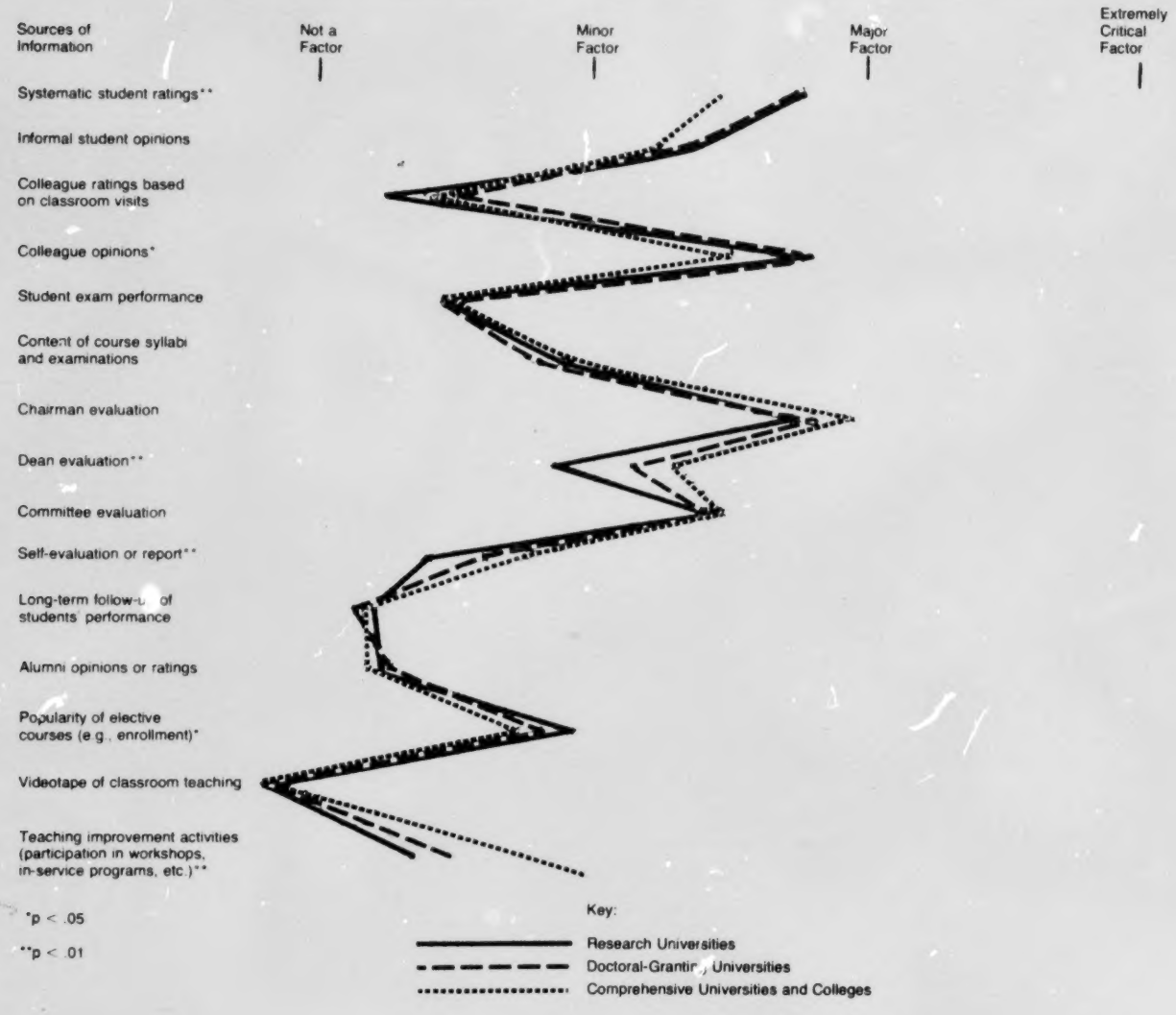
Table 4 presents a ranking of the current and preferred criteria by level of university. Respondents in Research Universities thought that quality of publications followed by classroom teaching should be the top two criteria, as indeed has been the case in current practice. However, they would put less emphasis on the number of publications and more on research or creative activity independent of publication and on student advising, including the supervision of student research. Respondents in the second and third level universities, where classroom teaching was ranked first as a current and preferred criterion, would also like to see student advising and supervision of student research emphasized.

Evaluating Teaching Performance

A summary of the responses to 15 possible sources of information used to evaluate the teaching performance of individual faculty members appears in Figures 3 and 4 and in Tables 5, 6, and 7.

Chairmen evaluations, systematic student ratings, and colleagues' opinions were generally the most influential indicators of teaching performance. Least used were videotapes of classroom teaching, the long-term follow-up of students, alumni ratings, colleague ratings based on classroom visits, and student examination performance. The weight given to several of the indicators differed by level of university (Figure 3). Systematic student ratings, for example, were less important at Comprehensive Universities and Colleges than at Research or Doctoral-Granting Universities. Deans' evaluations and self-evaluations were least important at Research Universities. Teaching improvement activities, such as participation in in-service programs or faculty development practices, tended to carry

Figure 3. Average weight given to various sources of information for evaluating teaching performance, by university level



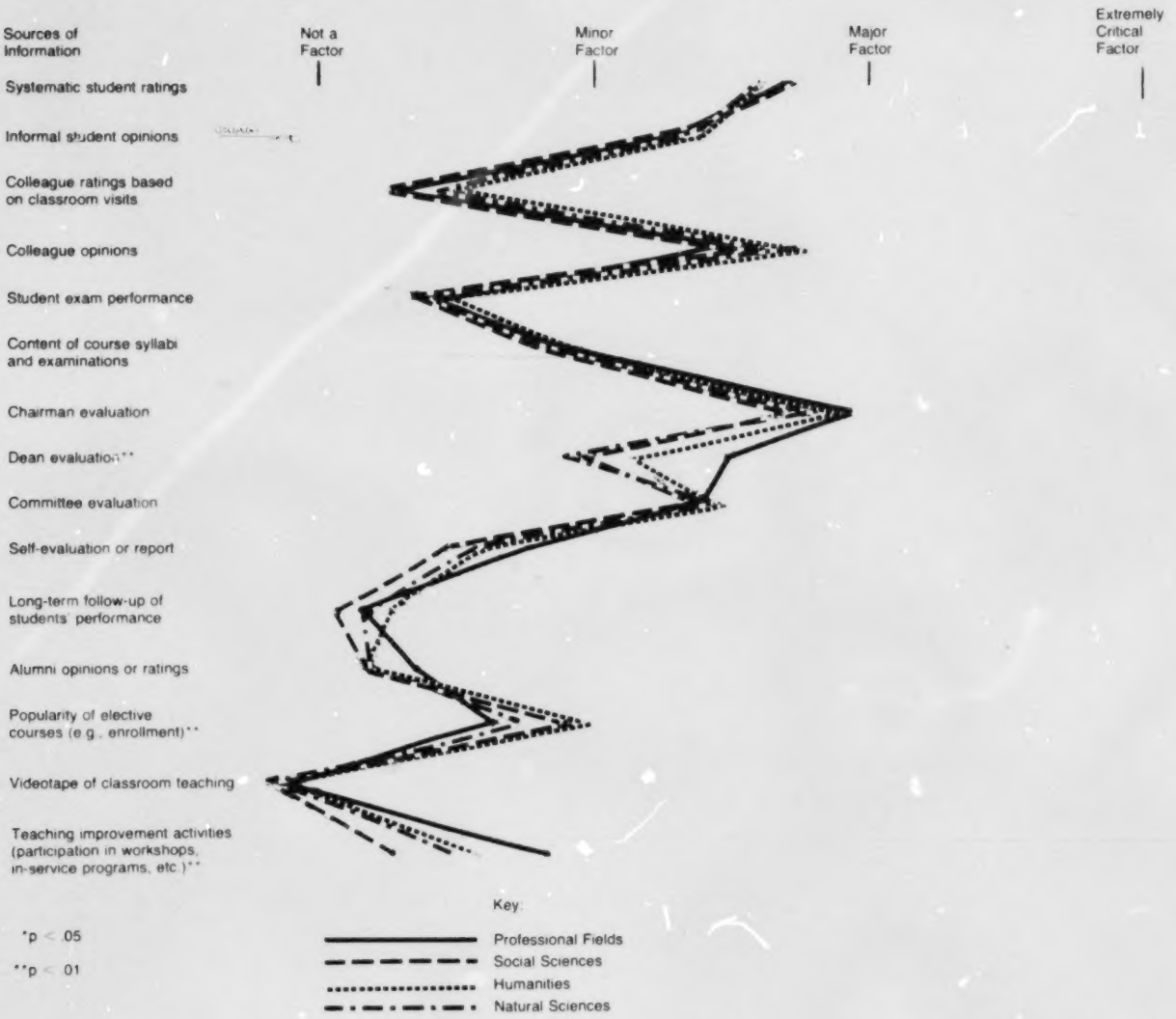
more weight in promotions at the Comprehensive Universities and Colleges than at the other two levels.

The department subgroups differed on only three of the teaching indicators (Figure 4). According to the chairmen, the professional departments emphasized deans' evaluations and teaching improvement activities more than did the other three department groups. Teaching improvement activities were given the least weight in social science departments. Both social science and humanities departments put slightly more emphasis on enrollment in elective courses as an indicator of teaching performance than did the professional or natural science departments.

Within each of the department subgroups there were also considerable differences in the emphasis given to several of the indicators (Table 5). For example, 21 percent of the social science departments reported systematic student ratings as an extremely critical factor, whereas 11 percent said they were not a factor. Committee evaluation varied even more, with almost equally high numbers of departments including it as an extremely critical factor or not a factor at all.

Tables 5, 6, and 7 report the preferences of department heads in evaluating teaching. Changes in the absolute values of current vs. preferred weights were also inspected.

Figure 4. Average weight given to various sources of information for evaluating teaching performance, by departmental subgroup



Although the respondents said that systematic student ratings should have slightly more importance than they currently do, few would make them a critical factor (Tables 5 and 7). Colleagues' opinions should be less important, but colleague ratings of teaching based on classroom visits, currently not a factor in many departments, should be given much greater emphasis. The content of course syllabi and of examinations should also be more important, moving from seventh- or eighth-ranked in current emphasis to fifth-ranked in preferred importance. The popularity of elective courses, respondents generally agreed, should receive relatively less importance, although they preferred to give more emphasis to long-term follow-up of student performance (Tables 6 and 7). Several sources of information, the respondents said, should have more importance in evaluating teaching, although there was little change in their relative position as indicated by the rankings. These were: student examination performance, alumni opinions or ratings, teaching improvement activities, and videotapes of classroom teaching. Apparently department heads would like to attach more importance to a wider variety of information than they now do. Whether it is feasible to do so with some of the criteria—for example, alumni ratings of teachers—is another question.

Table 5. Current and Preferred Sources of Information for Evaluating Teaching Performance

	Percentage Indicating Extremely Critical Factor and Not a Factor															
	Current Use and Importance								Importance Each *Should* Have							
	Professional Fields N=116		Social Sciences N=102		Humanities N=104		Natural Sciences N=131		Professional Fields N=116		Social Sciences N=102		Humanities N=104		Natural Sciences N=131	
	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.
Systematic student ratings	15	5	21	11	24	7	21	9	16	1	20	8	23	3	15	4
Informal student opinions	1	6	4	7	1	8	2	5	1	9	1	14	1	10	3	14
Colleague ratings based on classroom visits	2	50	4	54	9	47	5	34	9	19	10	39	11	22	8	19
Colleague opinions	12	9	10	8	17	6	8	3	9	4	5	11	13	7	7	7
Student examination performance	4	31	4	38	4	22	2	32	2	38	1	38	5	31	2	32
Content of course syllabi and examination	2	38	1	38	5	31	2	32	6	8	2	15	11	13	5	14
Chairman evaluation	23	5	16	10	24	8	17	5	25	2	13	10	24	7	14	5
Dean evaluation	22	12	10	28	13	27	7	24	15	11	3	30	9	25	4	26
Committee evaluation	22	20	24	25	27	18	24	16	18	12	21	19	24	13	22	11
Self-evaluation or report	3	37	1	45	4	40	3	35	6	19	2	38	6	21	5	27
Long-term follow-up of students' performance	2	59	1	61	6	51	2	47	6	23	5	38	10	20	7	20
Alumni opinions or ratings	1	48	1	57	0	57	0	45	3	34	0	38	4	32	2	31
Popularity of elective courses (e.g., enrollment)	1	43	2	25	1	23	1	40	0	34	1	25	0	27	0	38
Videotape of classroom teaching	0	62	0	64	1	71	1	61	4	47	2	50	1	51	2	56
Teaching improvement activities (participation in workshops, in-service programs, etc.)	0	29	1	49	4	48	1	45	5	14	1	31	4	22	1	27

Table 6. Ranking of Current and Preferred Sources of Information for Evaluating Teaching Performance, by Departmental Subgroup[1]

	Current Use and Importance				Importance Each *Should* Have			
	P	S	H	N	P	S	H	N
Systematic student ratings	2	1	3	3	2	1	2	2
Informal student opinions	6	5	5	5	9	6	6	5
Colleague ratings based on classroom visits	13	11	11	11.5	8	7	7	7
Colleague opinions	3	3	2	2	4	4	4	4
Student examination performance	11	10	12	10	12	11	12	11.5
Content of course syllabi and examination	7	8	8	6	5	5	5	6
Chairman evaluation	1	2	1	1	1	2	1	1
Dean evaluation	4	6	6	7	6	8	8	9.5
Committee evaluation	5	4	4	4	3	3	3	3
Self-evaluation or report	9	9	9	9	10	12	10	11.5
Long-term follow-up of students' performance	14	14	13	13	11	10	9	8
Alumni opinions or ratings	12	13	14	14	13	13	14	13
Popularity of elective courses (e.g., enrollment)	10	7	7	8	14	9	13	14
Videotape of classroom teaching	15	15	15	15	15	15	15	15
Teaching improvement activities (participation in workshops, in-service programs, etc.)	8	12	10	11.5	7	14	11	9.5

P = Professional Fields: N = 116 (rho = .86)
S = Social Sciences: N = 102 (rho = .88)
H = Humanities: N = 104 (rho = .84)
N = Natural Sciences: N = 131 (rho = .81)

Rank difference correlations between current importance and preferred importance are shown in parentheses.

Table 7. Ranking of Current and Preferred Sources of Information for Evaluating Teaching Performance, by University Level[1]

	Current Use and Importance				Importance Each *Should* Have			
	All	I	II	III	All	I	II	III
Systematic student ratings	2.5	2	2	4	1	1	1	2
Informal student opinions	5	5	5	6	6	6	7	7.5
Colleague ratings based on classroom visits	12	12	10	12	7	7	6	9
Colleague opinions	2.5	1	3	2	3	3	3	4
Student examination performance	11	9	12	11	12	9	13	12
Content of course syllabi and examinations	7	7	8	7	5	5	5	5
Chairman evaluation	1	3	1	1	2	2	2	1
Dean evaluation	6	8	6	5	8	11	9	6
Committee evaluation	4	4	4	3	4	4	4	3
Self-evaluation or report	9	10	9	8.5	11	14	10	10
Long-term follow-up of students' performance	14	14	14	13	9	8	8	11
Alumni opinions or ratings	13	13	13	14	14	12.5	12	14
Popularity of elective courses (e.g., enrollment)	8	6	7	10	13	10	14	13
Videotape of classroom teaching	15	15	15	15	15	15	15	15
Teaching improvement activities (participation in workshops, in-service programs, etc.)	10	11	11	8.5	10	12.5	11	7.5

Departments grouped by university classification as follows:
I = Research Universities: N = 158 (rho = .79)
II = Doctoral-Granting Universities: N = 122 (rho = .77)
III = Comprehensive Universities and Colleges: N = 173 (rho = .93)

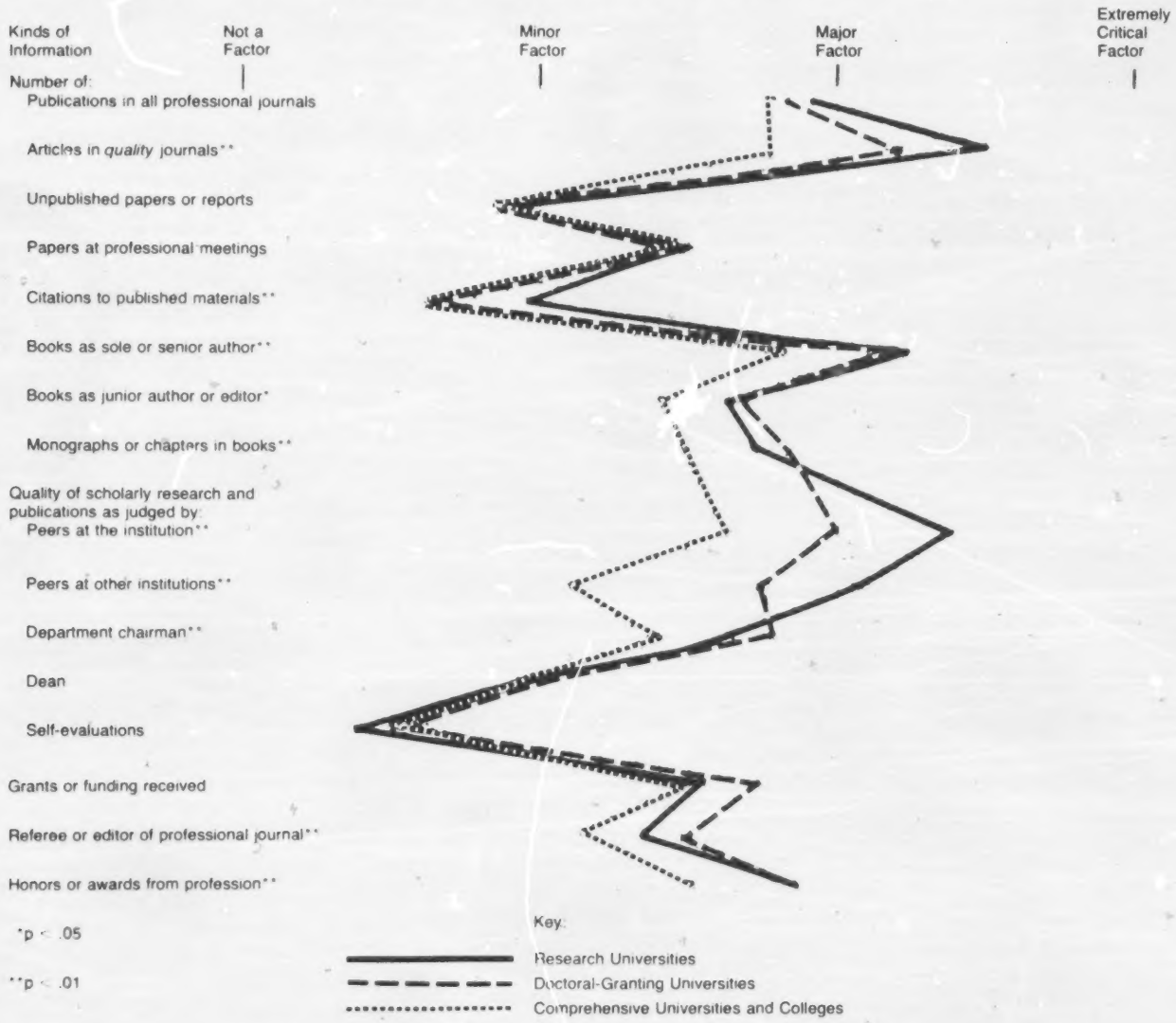
Rank difference correlations between current importance and preferred importance are shown in parentheses.

Evaluating Scholarship and Research

Figures 5, 6, 7, and 8 and Tables 8, 9, and 10 summarize the responses of department heads to the kinds of information used in evaluating scholarship or research performance. In general, the number of articles published in quality journals, the number of books of which the faculty member is the sole or senior author, and the quality of one's research and publications as judged by peers at the institution are the three most important types of information used. But there are important variations between departmental groups and university levels. The Research Universities tend to put more emphasis on the judgment of peers—both from within and outside the institution—than do either Doctoral-Granting Universities or Comprehensive Universities and Colleges (Figure 5). Comprehensive Universities and Colleges put the least emphasis on most of the criteria listed, with the Doctoral-Granting Universities falling between them and the Research Universities.

The departmental groups also differed in how they evaluated research and scholarship (Figure 6). The professional departments, compared with those in the arts and sciences, placed less emphasis on peer judgments and on the number of quality publications. Essentially, however, the way in which departments evaluated research interacted with the level of the university—that is, the departments differed according to the level of the university. This interaction can be illustrated by use of a canonical discriminant function analysis of the 12 groups (the four departments within each of the three levels of universities) and of the 16 kinds of information used to assess research performance. The analysis provides an optimal discrimina-

Figure 5. Average weight given to various kinds of information used to evaluate scholarship or research performance, by university level



tion between the 12 groups on the basis of the 16 kinds of information, using the mean responses of department heads as the unit of analysis.

The results of the canonical discriminant function analysis, which are summarized in Appendix B and in Figures 7 and 8, indicate two major dimensions of variation among the 12 groups on the 16 kinds of information. Each of these two dimensions accounted for about a third of the variance among the 12 groups for a total of 67 percent (Appendix B). The group centroids are plotted on the discriminant axes in Figure 7.[2] For example, in the figure, I P identifies the position of professional departments in Research Universities, II S represents social science departments in Doctoral-Granting Universities, and so forth. In addition to these group centroids, Figure 7 includes 16 lines of vectors originating from the center of the plot. These 16 lines represent the way the 16 kinds of information used to evaluate research and scholarship project into the discriminant space. The important features of these vectors are their directions and their relative lengths. By visualizing the direct perpendicular projection of the 12 group centroids onto these vectors, one can get a feeling for the relative ordering among the department university level groups as to how they evaluate research performance.

A simplified illustration of the data appears in Figure 8. The two axes were rotated by visually estimating where they best fit the canonical variates. The descriptions at the

[2] The discriminant function weights given in Appendix B are difficult to interpret. Therefore, in order to better understand the results, the between-groups correlations between the original responses and the discriminant axes were studied. These correlations have the advantage of being treated much like factor loadings of observed variables on orthogonal factor axes—the canonical discriminant functions (Cliff and Krus, 1976).

Figure 6. Average weight given to various kinds of information used to evaluate scholarship or research performance, by departmental subgroup

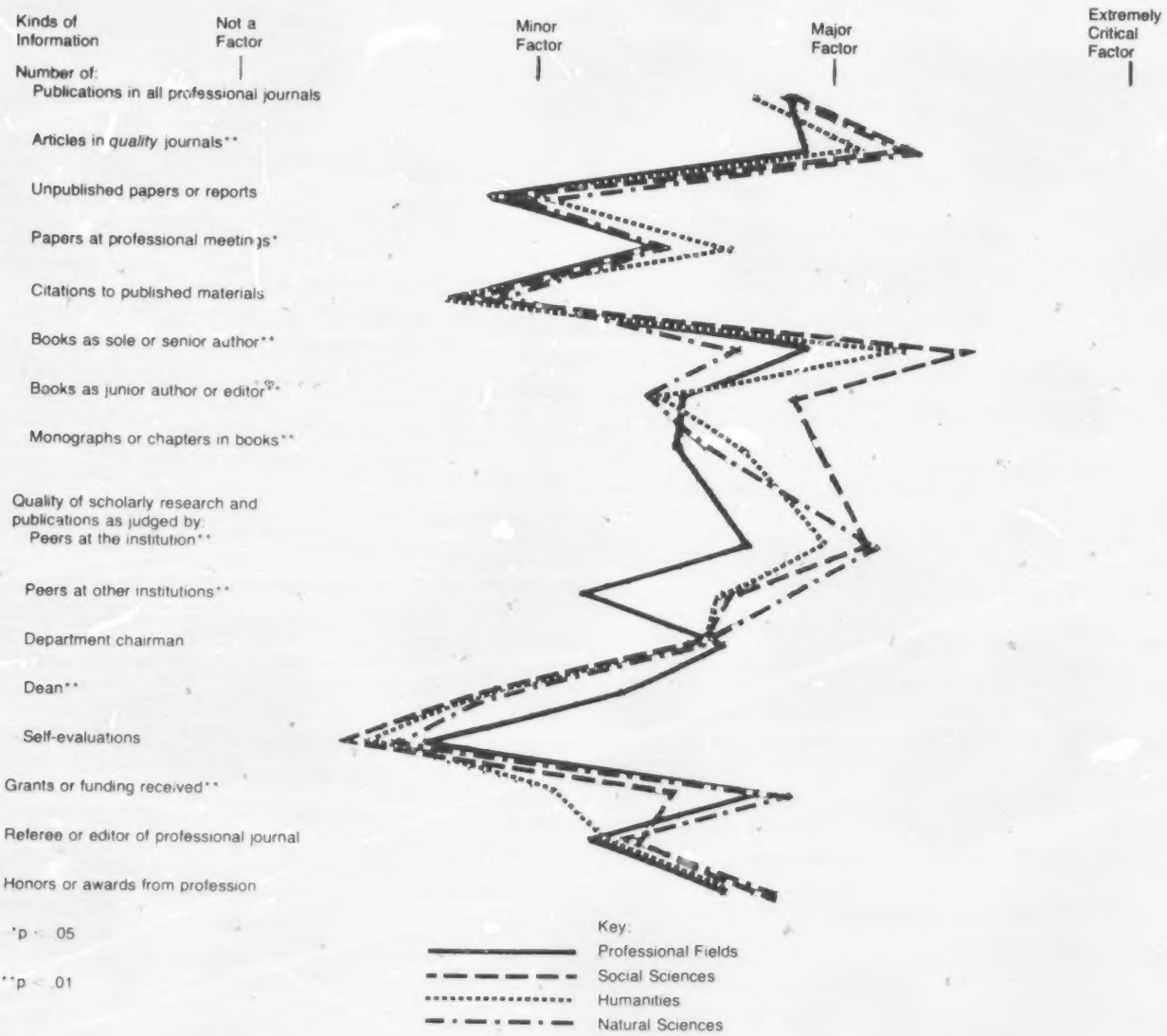


Figure 7. Plot of university level and departmental subgroup centroids and the 16 kinds of information used to evaluate scholarship and research in the space defined by the two largest canonical discriminant axes

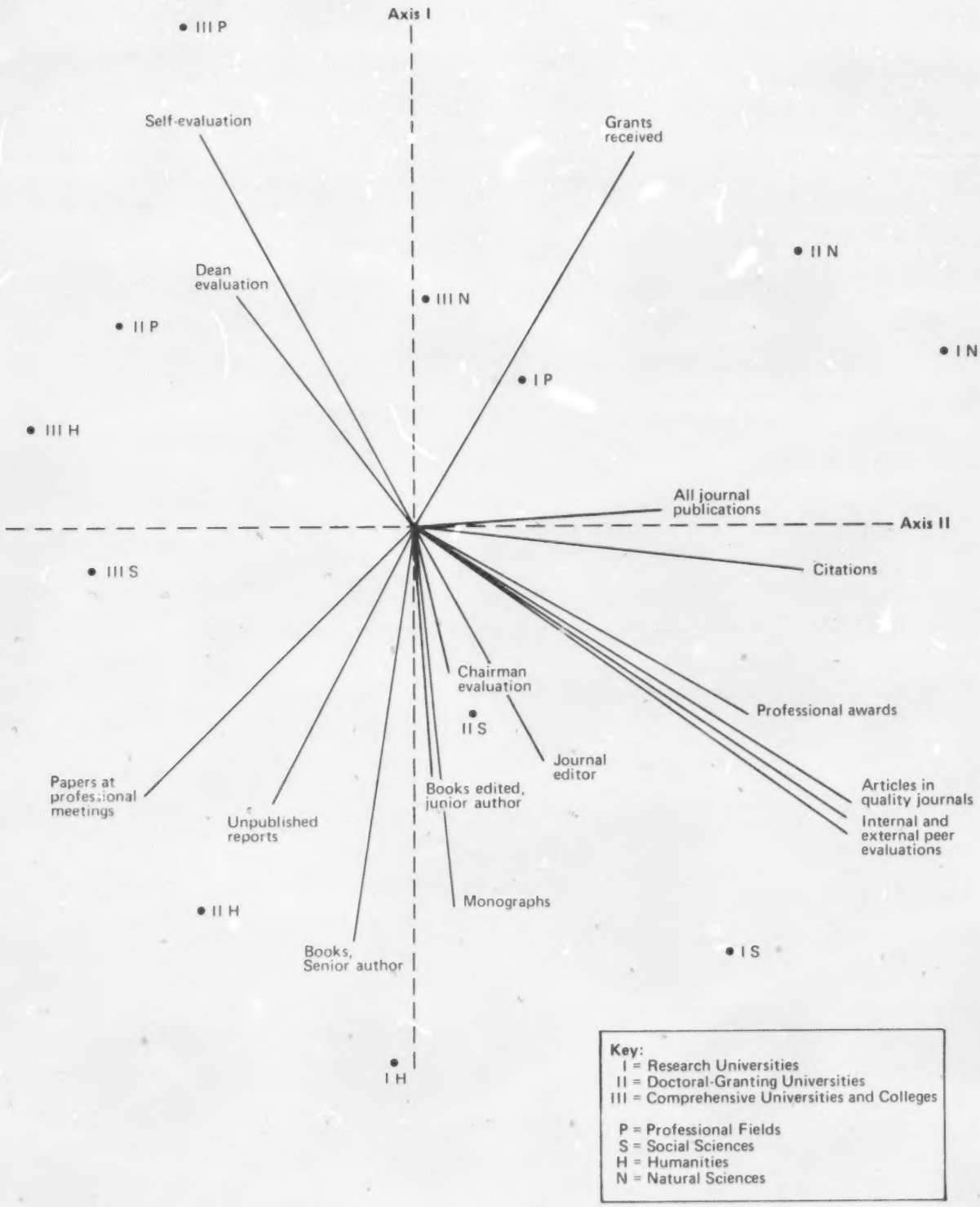


Figure 8. Rotated canonical discriminant axes, kinds of information used to evaluate scholarship and research performance

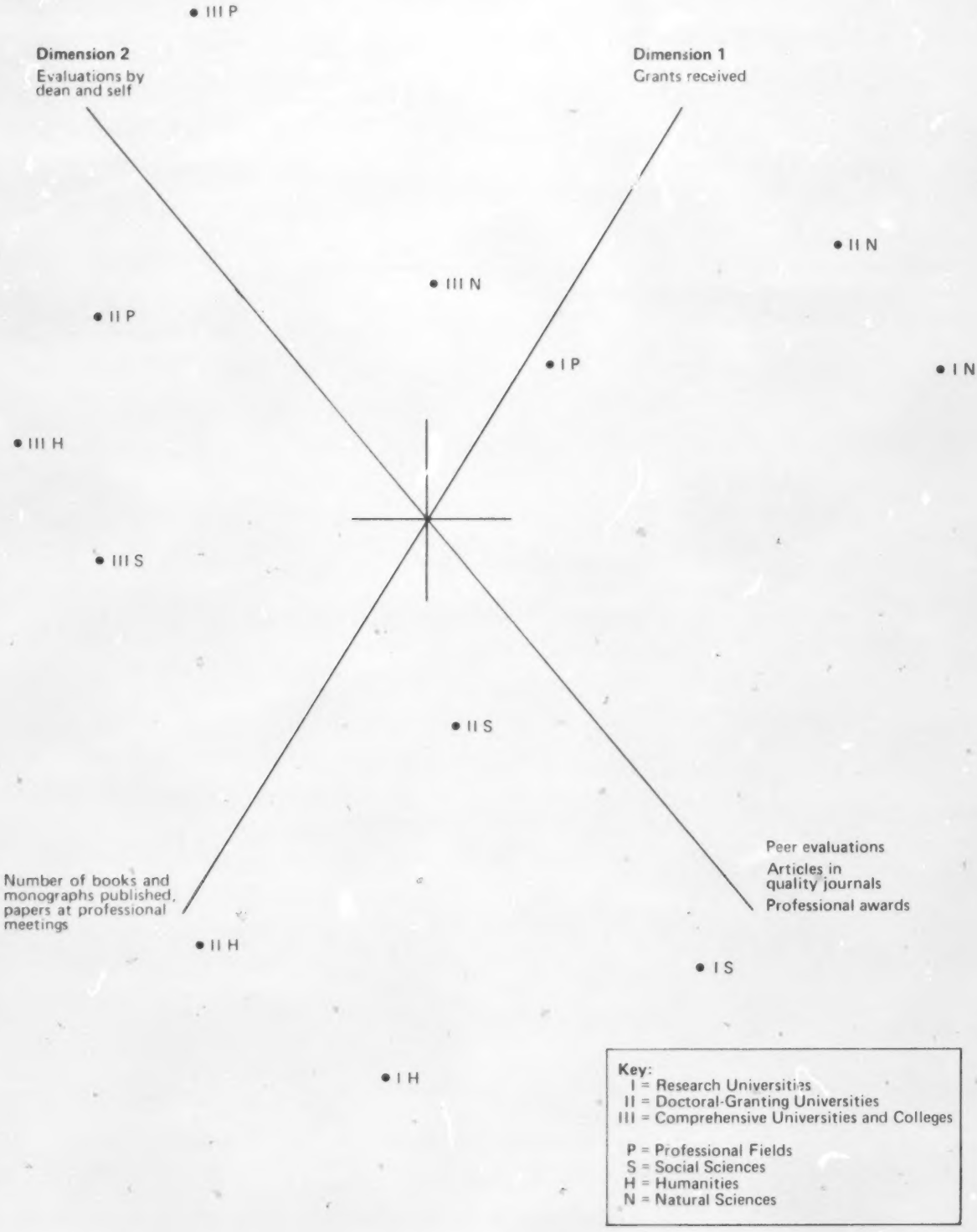


Table 8. Information Used to Evaluate Scholarship or Research Performance

	Percentage Indicating Extremely Critical Factor and Not a Factor															
	Current Use and Importance								Importance Each *Should* Have							
	Professional Fields N=116		Social Sciences N=102		Humanities N=104		Natural Sciences N=131		Professional Fields N=116		Social Sciences N=102		Humanities N=104		Natural Sciences N=131	
	Ext. Crit.	Not A Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.	Ext. Crit.	Not a Fact.
Number of:																
Publications in all professional journals	21	4	14	2	14	6	18	1	11	3	14	2	13	3	15	1
Articles in *quality* journals	29	3	46	2	37	3	40	1	44	1	57	1	45	1	45	1
Unpublished papers or reports	0	23	1	24	1	18	0	22	1	17	2	16	1	12	1	23
Papers at professional meetings	3	6	4	3	13	3	3	2	3	2	4	5	13	1	5	2
Citations to published materials	5	36	2	32	1	33	4	28	7	15	5	19	4	12	8	11
Books as sole or senior author	25	6	53	1	54	4	15	5	26	2	53	1	59	2	15	1
Books as junior author or editor	5	8	14	1	10	6	5	6	5	3	14	1	8	2	6	3
Monographs or chapters in books	4	6	18	0	17	4	6	4	7	2	20	1	20	2	7	2
Quality of scholarly research and publications as judged by:																
Peers at the institution	15	10	36	7	36	6	36	3	20	3	35	6	37	1	37	2
Peers at other institutions	16	22	25	14	31	15	36	10	19	5	32	5	38	2	38	2
Department chairman	14	7	16	14	19	12	10	6	12	7	17	12	15	7	12	8
Dean	13	15	8	32	5	29	2	27	8	15	5	36	3	26	1	26
Self-evaluations	4	43	1	46	2	51	2	42	7	30	2	51	4	40	2	38
Grants or funding received	20	8	7	11	6	21	16	1	13	4	10	11	4	11	10	3
Referee or editor of professional journal	0	12	7	8	7	8	5	7	4	4	9	2	10	4	3	4
Honors or awards from profession	14	6	16	2	12	3	18	2	21	0	20	1	14	2	18	0

end of each axis or dimension summarize which of the 16 variables seem to best define the pole. The first dimension contrasts grants received with the number of books published and papers presented at professional meetings. As indicated by their position furthest out from the center, natural science departments in the first two levels of universities (Research and Doctoral-Granting Universities), followed at some distance by third-level natural science departments (Comprehensive Universities and Colleges) and first-level professional departments, were at the "grants received" end of the continuum. As Figures 7 and 8 further illustrate, the humanities departments in the top two levels of institutions were at the "books-papers" pole. The second dimension contrasts evaluations by deans and by the individual faculty member (self-evaluation) with peer and professional evaluations. The latter include judgments of research or publications by peers from inside and outside the institution, articles published in quality journals, awards from the profession, and, to a lesser extent, the number of citations to the individual's published materials and journal editorship. Evaluations by deans and self-evaluations were emphasized most in professional departments at second- and third-level institutions as well as in humanities departments in third-level universities. Peer and professional evaluations (for example, articles in quality journals) were emphasized most in social science departments in first-level universities. To some extent, humanities and natural science departments in first-level universities also emphasized peer and professional evaluations. In sum, the assessment of research and scholarship, as Figures 7 and 8 help to make clear, varies considerably depending on the type of department and level of institution.

The current and preferred importance of each criterion for assessing research and scholarship is summarized in Tables 8, 9, and 10. Table 8 also includes the percentages of departments that use the criterion as an extremely critical factor or not at all. The quality of publications as judged by peers, department heads, or deans was extremely important in many departments, yet was not used as a criterion in many others. For example, 25 percent of the respondents in the social science departments reported peers at other institutions to be critical in judging scholarship, while 14 percent did not call on these peers as judges at all.

Department heads in the Research Universities were generally satisfied with the relative and absolute weights they use in evaluating research or scholarship. Those in the

Table 9. Ranking of Current and Preferred Criteria for Evaluating Research or Scholarship Performance, by Departmental Subgroup[1]

	Current Use and Importance				Importance Each *Should* Have			
	P	S	H	N	P	S	H	N
Number of:								
Publications in all professional journals	3	5	4	5	8.5	8	8	6
Articles in *quality* journals	2	2	2	1	1	1	2	1
Unpublished papers or reports	14	13	13	14.5	15	14	14	15
Papers at professional meetings	10	11	6.5	11	11	12	7	11
Citations to published materials	15	15	15	13	14	13	15.5	13
Books as sole or senior author	1	1	1	7	2	2	1	5
Books as junior author or editor	8	6	10	10	10	7	11	10
Monographs or chapters in books	9	4	5	8	6	5	5	8
Quality of scholarly research and publications as judged by:								
Peers at the institution	5.5	3	3	2	3	3	3	2
Peers at other institutions	12	8	8	3	5	4	4	3
Department chairman	7	9	9	9	8.5	9	9	9
Dean	11	14	14	14.5	13	15	15.5	14
Self-evaluations	16	16	16	16	16	16	10	16
Grants or funding received	4	10	12	4	7	10	13	7
Referee or editor of professional journal	13	12	11	12	12	11	12	12
Honors or awards from profession	5.5	7	6.5	6	4	6	6	4

[1] P = Professional Fields; N = 116 (rho = .82)
S = Social Sciences; N = 102 (rho = .94)
H = Humanities; N = 104 (rho = .89)
N = Natural Sciences; N = 131 (rho = .97)

Rank difference correlations between current importance and preferred importance are shown in parentheses.

Table 10. Ranking of Current and Preferred Criteria for Evaluating Research or Scholarship Performance, by University Level[1]

	Current Use and Importance				Importance Each *Should* Have			
	All	I	II	III	All	I	II	III
Number of:								
Publications in all professional journals	4	5	5.5	2	7	7	11	7
Articles in *quality* journals	1	1	1	3	1	1	1	1
Unpublished papers or reports	14	15	14	14	15	15	15	13
Papers at professional meetings	11	11	12	7.5	11	9.5	12	9
Citations to published materials	15	13	15	15.5	13	13	14	14
Books as sole or senior author	2	3	2	1	2	4	2	2
Books as junior author or editor	10	9	10	9	9	9.5	9	8
Monographs or chapters in books	6	7	4	7.5	6	6	6	6
Quality of scholarly research and publications as judged by:								
Peers at the institution	3	2	3	4	3	2	3	3
Peers at other institutions	7	4	8	12	4	3	4	5
Department chairman	8.5	8	7	10	8	8	7	12
Dean	13	14	13	13	14	14	13	15
Self-evaluations	16	16	16	15.5	16	16	16	16
Grants or funding received	8.5	10	9	6	10	11	8	10
Referee or editor of professional journal	12	12	11	11	12	12	10	11
Honors or awards from profession	5	6	5.5	5	5	5	5	4

[1] Departments grouped by university classification as follows:
I = Research Universities; N = 158 (rho = .98)
II = Doctoral-Granting Universities; N = 122 (rho = .92)
III = Comprehensive Universities and Colleges; N = 173 (rho = .83)

Rank difference correlations between current importance and preferred importance are shown in parentheses.

Comprehensive Universities and Colleges thought that more weight should be given to judgments made by peers at other institutions and in the number of publications in quality journals. Most respondents, regardless of level of university or department, preferred to put relatively less emphasis on the sheer number of publications in all types of journals. The natural science and professional departments also preferred to put slightly less weight on grants received than they currently do.

Discussion

Teaching, research, and service are usually enumerated as the three functions of most universities and, therefore, presumably the three major concerns of faculty members. But the responses given by 453 department heads from 134 universities indicate that public or university service is generally given little importance in evaluating faculty members for decisions regarding tenure, salary, and promotion. The Research Universities, with large Ph.D. programs and heavy financial support for research, emphasize research, as might be expected. Teaching, though, according to the respondents, is a close second in importance. At the Doctoral-Granting Universities and the Comprehensive Universities and Colleges, teaching was ranked first in importance, followed closely by research. Personal qualifications, such as having an advanced degree and the appropriate experience, were the only other criteria considered major factors in advancement decisions. Therefore, in addition to community and public service, student advising and service to the institution were also seen as relatively minor factors. In short, this study suggests that once the appropriate credentials are in hand, most universities and departments evaluate most faculty members as researchers-scholars and classroom teachers, with varying emphasis depending on the level of the university or the type of the department. Of course, it is important to note that the chairmen were responding to the criteria as they applied to their entire department; several pointed out that individual

faculty members in the department frequently have varying responsibilities and that the importance of any criterion depends on what those responsibilities are. Some chairmen also added that the significance of the various criteria depends on the level of the advancement in question and that each candidate was expected to be outstanding in at least one area (usually research or teaching).

Whether teaching does, in fact, receive as much weight in promotion and tenure decisions as chairmen report or whether they are merely paying lip service to its importance is difficult to know. There is reason to doubt: The faculty in the previously reported Oregon State study (Thorne, Scott, and Beaird, 1976) viewed publications as the most influential piece of evidence in decisions on advancement. It may be that research activity would have ranked ahead of classroom teaching for all levels of universities (not just the Research Universities) if the questionnaire had listed a single criterion (as with classroom teaching) instead of three separate kinds of information (number of publications, quality of publications, and research and/or creative activity independent of publication). This may have diffused the emphasis that research and scholarship activity actually receives.

The results of this study also suggest an increase in the use of some of the more systematic and tangible kinds of evidence for evaluating teaching performance. In comparison with the Astin and Lee survey a decade earlier, systematic student ratings and, to some extent, the content of course syllabi and examinations are used more frequently. Other recent studies, as discussed earlier, have noted a similar increase in the use of student ratings for tenure and promotion decisions. Surprisingly, according to the responses of department heads in this study, student ratings were even more important in the Research Universities and the Doctoral-Granting Universities than in the Comprehensive Universities and Colleges. Possibly the larger size of the Research and Doctoral-Granting Universities fostered the use of formal student ratings as evidence of teaching effectiveness.

Although there appears to be an increase in the use of at least some substantive evidence for evaluating teaching, apparently its use falls far short of what department heads think is needed. They would like to see even more emphasis on systematic student ratings and an evaluation of the content of course syllabi and examinations. They also believe that formal colleague ratings based on classroom visits and a long-term follow-up of how students perform should get more weight than they currently do. But both formal colleague ratings and the assessment of long-term student outcomes are difficult measures to obtain in any systematic and usable way. Colleague ratings based primarily on classroom observation would in most instances not be reliable enough to use in making decisions on tenure and promotion—at least not without faculty members investing considerable time in visitations or training sessions (Centra, 1975). And the assessment of student learning years after the students have completed a course, while possibly an ideal indicator of the long-term effects of a course or teacher, is unwieldy to administer; moreover, the effects are difficult to attribute to a specific teacher. An assessment of end-of-course learning, although more manageable, is also subject to misinterpretation when used in tenure and promotion decisions (Centra, 1977). No doubt, however, student learning is an important criterion in assessing teaching, and it may be that colleagues (or a department committee) could help review and judge pre-post achievement gains, student projects, or other evidence of course outcomes.

In comparison with the evaluation of teaching, the evaluation of research makes possible the use of much more tangible evidence, although judgments still play an important role in the evaluation. Articles, books, awards from the profession, and, in the natural sciences, grants received are the most important kinds of information used. Quality of research, according to the chairmen, is and should be the primary consideration. Therefore, in most academic fields the number of articles in quality journals is of prime importance, along with the judgments of research or scholarship made by peers at the institution. In fact, chairmen feel that more use should be made of peers at other institutions as additional reviewers of the quality of a faculty member's research or scholarship.

But in general, the evaluation of research and scholarship depends very much on the type of department and the level of institution. The number of books and papers produced, for example, is especially important in humanities departments at the top two levels of universities, but not at the third level. Peer judgments of research and the number of articles in quality journals are important in social science departments in the Research Universities but not in social science departments in the Comprehensive Universities and Colleges. These variations point out that the criteria used to evaluate research and scholarship should be set at the department level rather than at the institutional level—or, for that matter, by the discipline as a whole. Indeed, the same might be said for evaluating total faculty performance: the individual departments probably need some flexibility in setting their own criteria and standards of performance. Department procedures should be considered within the general framework provided by university policies and, as department heads in this study indicated, should be based on as much hard evidence as possible.

References

Astin, A. W., & Lee, C. B. T. Current practices in the evaluation and training of college teachers. In C. B. T. Lee (Ed.), *Improving college teaching*. Washington, D.C.: American Council on Education, 1967, 296–311.

Bejar, I. I. A survey of selected administrative practices supporting student evaluation of instruction programs. *Research In Higher Education*, 1975, 3, 77–86.

Boyd, J. E., & Schietinger, E. F. *Faculty evaluation procedures in southern colleges and universities*. Atlanta, Ga.: Southern Regional Education Board, 1976.

Carnegie Commission on Higher Education. *A classification of institutions of higher education*. Berkeley, Calif., 1973.

Centra, J. A. Colleagues as raters of classroom instruction. *Journal of Higher Education*, 1975, 46, 327–337.

Centra, J. A. The how and why of evaluating teaching. In J. A. Centra (Ed.), *Renewing and evaluating teaching*. San Francisco. Jossey-Bass, 1977.

Cliff, N., & Krus, D.J. Interpretation of canonical analysis. Rotated vs. unrotated solution. *Psychometrika*, 1976, 41, 35–42.

Gustad, J. W. Policies and practices in faculty evaluation. *Educational Record*, 1961, 42, 194–211.

Seldin, P. *How colleges evaluate professors*. Croton-on Hudson, N.Y.: Blythe-Pennington, Ltd., 1975.

Thorne, G. L., Scott, C. S., & Beaird, J. H. Assessing faculty performance: Final project report. Monmouth, Oregon: Oregon State System of Higher Education, Teaching Research Division, 1976.

Appendix A

Graduate Record Examinations Board
PRINCETON, NEW JERSEY 08540 • AREA CODE 609 921-9000

IN AFFILIATION WITH
The Association of Graduate Schools
The Council of Graduate Schools

1975-1976

Sanford S. Elberg
University of California at Berkeley
Chairman

•

Richard Armitage
Ohio State University

May Brodbeck
University of Iowa

Beverly Cassara
Federal City College

David R. Deener
Tulane University

I. Wesley Elliott
Fisk University

Wytze Gorter
University of Hawaii at Manoa

Arthur F. Jackson
North Carolina Agricultural
and Technical State University

Lyle V. Jones
University of North Carolina
at Chapel Hill

Sterling M. McMurrin
University of Utah

J. Boyd Page
Council of Graduate Schools
(ex officio)

George J. Resnikoff
California State University
at Hayward

Lorene L. Rogers
University of Texas
at Austin

Ben Rothblatt
University of Chicago

Harry H. Sisler
University of Florida

Donald J. White
Boston College

W. Dexter Whitehead
University of Virginia
at Charlottesville

•

Maryann A. Lear
Secretary to the Board

May 1976

Dear Colleague:

We are undertaking a survey of faculty evaluation procedures and need your assistance. The purpose of the survey is to investigate what departments are doing and what they think ought to be done when evaluating the performance of faculty members.

We hope you will be able to complete the enclosed questionnaire and return it to your graduate dean as soon as possible. He or she will forward your responses to researchers at Educational Testing Service who will summarize the results. Questionnaires are being distributed to a sample of department chairmen at about 150 universities.

The questionnaire asks about the methods being used in your department to evaluate faculty for promotion, salary increase, or tenure. Please indicate the importance given to each criterion when evaluating performance, including both teaching and scholarship or research performance. We would also like to know how you think various sources of information ought to be used. We recognize that final decisions are frequently made beyond the department level; however, recommendations by departments are critical. Therefore, this questionnaire focuses on procedures that are used to arrive at departmental recommendations.

Because the response format may not allow you to adequately describe what is done at your department in faculty evaluation (or what you think should be done), we have also provided space for your additional comments. The survey report will not identify individuals or institutions by name.

Thank you for your help with the survey. A report of the results will be sent to you at your request and we trust it will be of interest to you.

Sincerely,

Sanford S. Elberg

Sanford S. Elberg
Chairman

cc: Maryann A. Lear

SURVEY OF DEPARTMENTAL PRACTICES IN EVALUATING
FACULTY PERFORMANCE

I. Listed below are criteria that might be taken into account in evaluating faculty members for promotion, salary increase, or tenure. We would like you to give your best judgment about:

A. how important each criterion is in personnel judgments currently being made about faculty members within the department; (Column A)

B. how important you think each criterion should be. That is, given the goals of your department and institution, how much weight ought to be placed on each factor? (Column B) Respond as follows:

1 = Not a factor
2 = Minor factor
3 = Major factor
4 = Extremely critical factor

0 = Not available

	A Current use and importance in personnel recommendations (Circle one response in each row)					B Importance each should have in personnel recommendations (Circle one response in each row)			
Classroom teaching	1	2	3	4	0	1	2	3	4
Number of publications	1	2	3	4	0	1	2	3	4
Quality of publications	1	2	3	4	0	1	2	3	4
Research and/or creative activity (independent of publication)	1	2	3	4	0	1	2	3	4
Supervision of student research, including serving on masters and doctoral committees	1	2	3	4	0	1	2	3	4
Student advising	1	2	3	4	0	1	2	3	4
Campus committee work, service to college	1	2	3	4	0	1	2	3	4
Activity in professional societies (hold office, edit journal, etc.)	1	2	3	4	0	1	2	3	4
Public or community service	1	2	3	4	0	1	2	3	4
Competing job offers	1	2	3	4	0	1	2	3	4
Consultation (government, business, etc.)	1	2	3	4	0	1	2	3	4
Personality factors	1	2	3	4	0	1	2	3	4
Personal qualifications (academic degrees, professional experience, etc.)	1	2	3	4	0	1	2	3	4
Other (specify)_____ _____	1	2	3	4	0	1	2	3	4

Comments about above criteria:

II. Listed below are sources of information that might be taken into account when evaluating the teaching performance of individual faculty members in connection with promotion, salary increases, or tenure recommendations. What is the current importance of each in your department and what importance do you think each should have?

1 = Not a factor
2 = Minor factor
3 = Major factor
4 = Extremely critical factor
0 = Not available

Indicators of teaching performance	Current use and importance in evaluating teaching performance (Circle one response in each row)					Importance each should have in evaluating teaching performance (Circle one response in each row)			
Systematic student ratings	1	2	3	4	0	1	2	3	4
Informal student opinions	1	2	3	4	0	1	2	3	4
Colleague ratings based on classroom visits	1	2	3	4	0	1	2	3	4
Colleagues' opinions	1	2	3	4	0	1	2	3	4
Student examination performance	1	2	3	4	0	1	2	3	4
Content of course syllabi and examinations	1	2	3	4	0	1	2	3	4
Chairman evaluation	1	2	3	4	0	1	2	3	4
Dean evaluation	1	2	3	4	0	1	2	3	4
Committee evaluation	1	2	3	4	0	1	2	3	4
Self-evaluation or report	1	2	3	4	0	1	2	3	4
Long term follow up of how students perform	1	2	3	4	0	1	2	3	4
Alumni opinions or ratings	1	2	3	4	0	1	2	3	4
Popularity of elective courses (e.g., enrollment)	1	2	3	4	0	1	2	3	4
Video-tape of classroom teaching	1	2	3	4	0	1	2	3	4
Teaching improvement activities (participation in workshops, in-service programs, etc.)	1	2	3	4	0	1	2	3	4
Others (specify)_____ _____	1	2	3	4	0	1	2	3	4

Comments about above indicators:

III. Listed below are kinds of information that might be used to evaluate the scholarship or research performance of individual faculty members. Which ones are now used in your department when making personnel decisions? Which ones do you think should be used?

1 = Not a factor
2 = Minor factor
3 = Major factor
4 = Extremely critical factor
0 = Not available

Kinds of information	Current use and importance in evaluating scholarship or research performance (Circle one response in each row)	Importance each should have in evaluating scholarship or research performance (Circle one response in each row)
Number of:		
Publications in all professional journals	1 2 3 4 0	1 2 3 4
Articles in quality journals	1 2 3 4 0	1 2 3 4
Unpublished papers or reports	1 2 3 4 0	1 2 3 4
Papers at professional meetings	1 2 3 4 0	1 2 3 4
Citations to published materials	1 2 3 4 0	1 2 3 4
Books as sole or senior author	1 2 3 4 0	1 2 3 4
Books as junior author or editor	1 2 3 4 0	1 2 3 4
Monographs or chapters in books	1 2 3 4 0	1 2 3 4
Quality of scholarly research and publications as judged by:		
Peers at the institution	1 2 3 4 0	1 2 3 4
Peers at other institutions	1 2 3 4 0	1 2 3 4
Department chairman	1 2 3 4 0	1 2 3 4
Dean	1 2 3 4 0	1 2 3 4
Self-evaluations	1 2 3 4 0	1 2 3 4
Grants or funding received	1 2 3 4 0	1 2 3 4
Referee or editor of professional journal	1 2 3 4 0	1 2 3 4
Honors or awards from profession	1 2 3 4 0	1 2 3 4
Others (specify)_____		
_____	1 2 3 4 0	1 2 3 4

Comments about above information:

Please use the space below to comment on practices in your department or at your institution that may not be adequately reflected in your responses thus far. Are there practices that you think should be followed instead?

Name_____

Title or Rank_____

Department_____ University_____

Address_____

Highest degree offered by the department:_____

Number of full-time faculty members in the department:_____

Check here if you would like a copy of the final results. _____

Someone in your graduate dean's office is coordinating the distribution and collection of these questionnaires. Please return your completed form to this person.

Thank you for your assistance.

John A. Centra
Mary Jo Clark
Project Directors
Educational Testing Service
Princeton, New Jersey 08540
(609) 921-9000

Appendix B

Canonical Discriminant Function Analysis of Information Used to Evaluate Scholarship and Research[1]

Discriminant Function	*Root*	*Cumulative % Trace*
1	.5130	35.5
2	.4492	66.5
3	.1548	77.2
4	.1050	84.5
5	.0729	89.6
6	.0591	93.6
7	.0283	95.6
8	.0280	97.5
9	.0233	99.1
10	.0084	99.7
11	.0039	100.0

Wilk's Lambda Criterion = .287 for F (176, 3909) = 3.25

Test for residual after removing:	*Chi-square*	*df*
First discriminant function	365.95	150
First two discriminant functions	203.09	126
First three discriminant functions	139.91	104
First four discriminant functions	96.07	84
First five discriminant functions	65.16	66
First six discriminant functions	39.94	50
First seven discriminant functions	27.68	36
First eight discriminant functions	15.55	24
First nine discriminant functions	5.42	14
First ten discriminant functions	1.72	6

Discriminant Function Weights Scaled for Unit Within-Groups Variance on Original Responses

	Discriminant Function		
Variable	*I*	*II*	*III*
Number of:			
Journal publications	.179	.028	.025
Articles in quality journals	.015	.413	.233
Unpublished papers/reports	−.099	−.096	−.157
Papers at professional meetings	−.137	−.224	−.376
Citations to materials	.003	.121	−.209
Books as sole or senior author	−.568	−.453	.265
Books as junior author or editor	.254	.222	.256
Monographs or chapters	−.111	−.037	.159
Quality of work judged by:			
Peers at the institution	−.201	.287	−.115
Peers at other institutions	−.117	.259	−.314
Department chairman	−.088	−.180	.201
Dean	.132	−.093	−.113
Self-evaluations	.156	−.152	.087
Grants or funding received	.451	.211	.334
Journal editor or referee	−.035	−.055	.054
Professional awards/honors	−.048	.064	.048

[1]The author is grateful to Allen Yates for his help with the canonical discriminant analysis.